

Article

Patterns of shared signatures of recent positive selection across human populations

Kelsey Elizabeth Johnson & Benjamin F. Voight

Nature Ecology & Evolution (2018)

doi:10.1038/s41559-018-0478-6

Evolutionary biology Evolutionary genetics

Population genetics

Received: 28 June 2017

Accepted: 11 January 2018

Published online: 19 February 2018

Abstract

Signatures of recent positive selection often overlap across human populations, but the question of how often these overlaps represent a single ancestral event remains unresolved. If a single selective event spread across many populations, the same sweeping haplotype should appear in each population and the selective pressure could be common across populations and environments. Identifying such shared selective events could identify genomic loci and human traits important in recent history across the globe. In addition, genomic annotations that recently became available could help attach these signatures to a potential gene and molecular phenotype selected across populations. Here, we present a catalogue of selective sweeps in humans, and identify those that overlap and share a sweeping haplotype. We connect these sweep overlaps with potential biological mechanisms at several loci, including potential new sites of adaptive introgression, the glycoporphin locus associated with malarial resistance and the alcohol dehydrogenase cluster associated with alcohol dependency.

Subscribe to *Nature Ecology & Evolution* for full access:

\$99

Subscribe

READCUBE ACCESS:

\$4.99

rent

\$20.00

buy

Buy/Rent now

Additional access options:

Already a subscriber? [Log in now](#) or [Register](#) for online access.

[Login via Athens](#) | [Login via Shibboleth](#) | [Use a document delivery service](#) | [Purchase a site license](#)

Additional information

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Tishkoff, S. A. et al. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* **39**, 31–40 (2007).
2. Kamberov, Y. G. et al. Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell* **152**, 691–702 (2013).
3. Fumagalli, M. et al. Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* **349**, 1343–1347 (2015).

R

4. Pickrell, J. K. et al. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res.* **19**, 826–837 (2009).

5. Coop, G. et al. The role of geography in human adaptation. *PLoS Genet.* **5**, e1000500 (2009).

6. Metspalu, M. et al. Shared and unique components of human population structure and genome-wide signals of positive selection in South Asia. *Am. J. Hum. Genet.* **89**, 731–744 (2011).

7. Liu, X. et al. Detecting and characterizing genomic signatures of positive selection in global populations. *Am. J. Hum. Genet.* **92**, 866–881 (2013).

8. Aguet, F. et al. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).

9. Vernot, B. & Akey, J. M. Resurrecting surviving neandertal lineages from modern human genomes. *Science* **343**, 1017–1021 (2014).

10. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

11. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).

12. Sabeti, P. C. et al. Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).

13. McVicker, G., Gordon, D., Davis, C. & Green, P. Widespread genomic signatures of natural selection in hominid evolution. *PLoS Genet.* **5**, e1000471 (2009).

14. Scheet, P. & Stephens, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* **78**, 629–644 (2006).

15. Baum, J., Ward, R. H. & Conway, D. J. Natural selection on the erythrocyte surface. *Mol. Biol. Evol.* **19**, 223–229 (2002).

16. Wang, H. Y., Tang, H., Shen, C. K. J. & Wu, C. I. Rapidly evolving genes in human. I. The glycophorins and their possible role in evading malaria parasites. *Mol. Biol. Evol.* **20**, 1795–1804 (2003).

17. Ko, W. Y. et al. Effects of natural selection and gene conversion on the evolution of human glycophorins coding for MNS blood polymorphisms in malaria-endemic African populations. *Am. J. Hum. Genet.* **88**, 741–754 (2011).

18. Leffler, E. M. et al. Resistance to malaria through structural variation of red blood cell invasion receptors. *Science* **356**, eaam6393 (2017).

19. Leffler, E. M. et al. Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science* **339**, 1578–1582 (2013).

20. Raj, T. et al. Common risk alleles for inflammatory diseases are targets of recent positive selection. *Am. J. Hum. Genet.* **92**, 517–529 (2013).

21. van der Zanden, L. F. M. et al. Common variants in DGKK are strongly associated with risk of hypospadias. *Nat. Genet.* **43**, 48–50 (2011).

22. Geller, F. et al. Genome-wide association analyses identify variants in developmental genes associated with hypospadias. *Nat. Genet.* **46**, 957–963 (2014).

23. Paré, G. et al. Novel associations of CPS1, MUT, NOX4, and DPEP1 with plasma Homocysteine in a healthy population a genome-wide evaluation of 13 974 participants in the women's genome health study. *Circ. Cardiovasc. Genet* **2**, 142–150 (2009).

24. van Meurs, J. B. J. et al. Common genetic loci influencing plasma homocysteine concentrations and their effect on risk of coronary artery disease. *Am. J. Clin. Nutr.* **98**, 668–676 (2013).

25. Huerta-Sánchez, E. et al. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194–197 (2014).

26. Racimo, F., Sankararaman, S., Nielsen, R. & Huerta-Sánchez, E. Evidence for archaic adaptive introgression in humans. *Nat. Rev. Genet.* **16**, 359–371 (2015).

27. Racimo, F., Marnetto, D. & Huerta-Sánchez, E. Signatures of archaic adaptive introgression in present-day human populations. *Mol. Biol. Evol.* **34**, 296–317 (2017).

28. Ding, Q., Hu, Y., Xu, S., Wang, J. & Jin, L. Neanderthal introgression at chromosome 3p21.31 was under positive natural selection in east asians. *Mol. Biol. Evol.* **31**, 683–695 (2014).

29. Sams, A. J. et al. Adaptively introgressed Neandertal haplotype at the OAS locus functionally impacts innate immune responses in humans. *Genome Biol.* **17**, 246 (2016).

30. Grossman, S. R. et al. Identifying recent adaptations in large-scale genomic data. *Cell* **152**, 703–713 (2013).

31. Enard, D., Messer, P. W. & Petrov, D. A. Genome-wide signals of positive selection in human evolution. *Genome Res.* **24**, 885–895 (2014).

32. Yu, F. et al. Population genomic analysis of 962 whole genome sequences of humans reveals natural selection in non-coding regions. *PLoS One* **10**, e0121644 (2015).

33. Kamburov, A., Wierling, C., Lehrach, H. & Herwig, R. ConsensusPathDB—a database for integrating human functional interaction networks. *Nucleic Acids Res.* **37**, D623–D628 (2009).

34. Han, Y. et al. Evidence of positive selection on a class I ADH locus. *Am. J. Hum. Genet.* **80**, 441–456 (2007).

35. Edenberg, H. J. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Res. Health* **30**, 5–13 (2007).

36. Li, D., Zhao, H. & Gelernter, J. Strong association of the alcohol dehydrogenase 1B gene (ADH1B) with alcohol dependence and alcohol-induced medical diseases. *Biol. Psychiatry* **70**, 504–512 (2011).

37. Galinsky, K. J. et al. Fast principal-component analysis reveals convergent evolution of ADH1B in Europe and East Asia. *Am. J. Hum. Genet.* **98**, 456–472 (2016).

38. Gelernter, J. et al. Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol. Psychiatry* **19**, 41–49 (2014).

39. Pashos, E. E. et al. Large, diverse population cohorts of hiPSCs and derived hepatocyte-like cells reveal functional genetic variation at blood lipid-associated loci. *Cell Stem Cell* **20**, 558–570.e10 (2017).

40. Lee, S. L., Chau, G. Y., Yao, C. T., Wu, C. W. & Yin, S. J. Functional assessment of human alcohol dehydrogenase family in ethanol metabolism: significance of first-pass metabolism. *Alcohol. Clin. Exp. Res.* **30**, 1132–1142 (2006).

41. Matsuo, K. et al. Alcohol dehydrogenase 2 His47Arg polymorphism influences drinking habit independently of aldehyde dehydrogenase 2 Glu487Lys polymorphism: analysis of 2,299 Japanese subjects. *Cancer Epidemiol. Biomark. Prev.* **15**, 1009–1013 (2006).

42. Haller, B. C. & Messer, P. W. SLiM 2: flexible, interactive forward genetic simulations. *Mol. Biol. Evol.* **34**, 230–240 (2017).

43. Terhorst, J., Kamm, J. A. & Song, Y. S. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat. Genet.* **49**, 303–309 (2016).

44. R Development Core Team R: *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2016).

45. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

46. Busing, F. M. T. A., Meijer, E. & Leeden van der, R. Delete-m jackknife for unequal m. *Stat. Comput.* **9**, 3–8 (1999).

47. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

48. Vernot, B. et al. Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* **352**, 235–239 (2016).

49. Simonti, C. N. et al. The phenotypic legacy of admixture between modern humans and Neandertals. *Science* **351**, 737–741 (2016).

50. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

Acknowledgements

We thank C. Brown and E. Leffler for helpful comments that improved the quality of the manuscript. This work was supported by grants from the National Institutes of Health (NIDDK R01DK101478 to B.F.V., T32GM008216 to K.E.J.) and a fellowship from the Alfred P. Sloan Foundation (BR2012-087 to B.F.V.).

Author information

Affiliations



*Genetics and Gene Regulation Program, Cell and Molecular Biology Graduate Group,
Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*
Kelsey Elizabeth Johnson

*Department of Systems Pharmacology and Translational Therapeutics, Perelman School of
Medicine, University of Pennsylvania, Philadelphia, PA, USA*
Benjamin F. Voight

*Department of Genetics, Perelman School of Medicine, University of Pennsylvania,
Philadelphia, PA, USA*
Benjamin F. Voight

*Institute for Translational Medicine and Therapeutics, Perelman School of Medicine,
University of Pennsylvania, Philadelphia, PA, USA*
Benjamin F. Voight

Contributions

K.E.J. and B.F.V. planned the study. K.E.J. assembled input data and performed the experiments. K.E.J. and B.F.V. interpreted the data and wrote the paper.

Competing interests

The authors declare no competing financial interests.

Corresponding author

Correspondence to Benjamin F. Voight.

Supplementary information

Supplementary Information - https://static-content.springer.com/esm/art%3A10.1038%2Fs41559-018-0478-6/MediaObjects/41559_2018_478_MOESM1_ESM.pdf
Supplementary Note, Supplementary Figures 1–11.



Life Sciences Reporting Summary - https://static-content.springer.com/esm/art%3A10.1038%2Fs41559-018-0478-6/MediaObjects/41559_2018_478_MOESM2_ESM.pdf

Supplementary Table 1 - https://static-content.springer.com/esm/art%3A10.1038%2Fs41559-018-0478-6/MediaObjects/41559_2018_478_MOESM3_ESM.xlsx
iHS standardization tables.

Supplementary Tables 2–9 - https://static-content.springer.com/esm/art%3A10.1038%2Fs41559-018-0478-6/MediaObjects/41559_2018_478_MOESM4_ESM.xlsx

Supplementary Table 2. 1,000 Genomes population information. Supplementary Table 3. Top iHS intervals for 26 populations. Supplementary Table 4. Estimated effective population sizes and sweep interval counts. Supplementary Table 5. Inferred shared sweep intervals. Supplementary Table 6. Observed and null distribution of shared sweep intervals for each population pair. Supplementary Table 7. Intersection of sweep tags with the GWAS catalogue. Supplementary Table 8. Introgressed Neanderthal haplotypes in LD with sweep tag variants. Supplementary Table 9. Association of a sweeping haplotype in YRI with alcohol dependence and alcohol dehydrogenase expression.